Title: Method and System for Determining Autonomous System Transit Volumes

Inventor(s):   1.   William Girard McCollom

2.   Joseph Ronald Hunt

3.   Alexander Lawrence Tudor

4.   Lance Anthony Tatman

5.   William E. Woodcock, IV

Method and System for Determining Autonomous System Transit Volumes

BACKGROUND OF THE INVENTION

Technical Field of the Invention

5   The present invention relates to determining autonomous system (AS) transit and

terminating volumes. More particularly, the present invention relates to calculating AS

transit and terminating volumes in internet or other communication systems using routing

information bases and data flow statistics for destination internet protocol addresses.

Embodiments of the present invention create tables of AS transit and terminating volumes

10   that permit network administrators to evaluate the suitability of transit providers and peers.

Description of Related Art

In a communication or a data communication system, routers maintain forwarding

tables that include a prefix (i.e., an IP address and mask), a next hop IP address, and other

routing parameters. The forwarding tables are generated via the border gateway protocol

15   (BGP) and other routing protocols. Information from which routers derive the forwarding

tables includes additional information about the potential path of the routed traffic, such as

the destination autonomous system (AS) number (known as the terminating AS) and a list of

intermediate AS numbers that the traffic traverses in order to reach the destination AS.

Internet service providers that use routers can use tools provided by router vendors to

20   analyze data traffic routed by the routers. The data traffic analysis can be based on counters

maintained by the routers. The counters can be aggregated into data flow counts, which are

totals of the number of bytes of data traffic observed between two internet protocol entities.

The aggregated data flow counts permit a determination to be made of how much traffic was

relayed via a particular protocol between any two locations. The router usually relays these

1

data flow counters to another system for storage and/or analysis. An example of such a

system is a CISCO router that has NETFLOW capabilities that are enabled and that streams

data flow information to another system. The system runs a process that stores and

aggregates the data flow for later analysis. The information provided by a NETFLOW

5   analysis merely provides data traffic volumes for a particular traffic destination. Users of the

NETFLOW analysis cannot determine, for example, the intermediate networks on which the

data traffic traveled. The NETFLOW users can only determine where the data traffic

terminated.

Many network operators would like to be able to determine which network providers

10   carry most of their data traffic, so that the network operators can either link directly to high

traffic providers for their data traffic termination or negotiate better contracts with the high

transit providers based on the observed data traffic volumes. However, existing NETFLOW-

based and other similar analyses do not provide this information to a user. A method and

system are needed that permit network administrators to determine on which intermediate

15   networks data traffic has traveled in order to evaluate actual and potential peers and transit

providers.

SUMMARY OF THE INVENTION

20   The present invention provides a system and method that permits calculation of

autonomous system (AS) transit and terminating data flow volumes. The AS transit and

terminating volumes permit network administrators to determine on which intermediate

networks data traffic has traveled in order to evaluate actual and potential peers and transit

providers. Routing information base data, including at least one prefix and at least one

25   selected AS path, is obtained from at least one router. The routing information base includes

one or more AS paths including one or more AS numbers through which data flows

2

corresponding to a particular prefix traverse. Data flow statistics sorted by destination address are obtained from the at least one router. The routing information base data and the data flow statistics are correlated according to ASs at which the data flows terminate or traverse. Thus, ASs can be evaluated in order to determine how much data traffic traverses

5   or terminates at the ASs.

A method of determining AS volume data includes collecting data flow statistics for at least one router and collecting routing information base data for each of the at least one router. The routing information base data and the data flow statistics are correlated, thereby yielding AS volume data.

10   A system for determining AS volume data includes a data flow collection node, a routing information base collection node, and a correlation node. The data flow collection node is adapted to collect data flow statistics from at least one router. The routing information base collection node is adapted to periodically collect a routing information base data from the at least one router. The correlation node is adapted to correlate the routing information base data and the data flow statistics and thereby yield AS volume data.

15   A method of generating autonomous system volume data includes detecting at least one first data flow having a first volume and directed toward a first destination address using a first selected autonomous path in a routing information base. For each autonomous system in the first selected autonomous system path, a counter is incremented by an amount

20   indicating the first volume.

A method of generating autonomous system volume data includes detecting at least one first data flow having a first volume and directed toward a first destination address. For each autonomous system in a first synthetic autonomous system path, a counter is incremented by an amount indicating the first volume.

A method of generating autonomous system volume data includes detecting at least

one first data flow having a first volume and directed toward a first destination address using

a first selected autonomous path in a routing information base. For a terminating autonomous

system in the first selected autonomous system path, a counter is incremented by an amount

5   indicating the first volume step of analyzing.

Embodiments of the present invention permit AS transit and terminating volumes to

be determined. The AS transit and terminating volumes thus determined permit network

administrators to know data traffic volumes that traverse or terminate at various ASs. The

AS volumes permit the network administrators to evaluate the ASs as peers or transit

10   providers. Furthermore, the present invention provides embodiments with other features and

advantages in addition to or instead of those discussed above. Many of these features and

advantages are apparent from the Description below with reference to the following

Drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the present invention can be achieved by reference

to the following Detailed Description when taken in conjunction with the accompanying

drawings, wherein:

20   FIG. 1 is a block diagram that illustrates an exemplary system in accordance with

principles of the present invention;

FIG. 2 is a flow diagram that illustrates exemplary operation of a correlation node N_C

116 in accordance with principles of the present invention; and

FIG. 3 is a flow diagram that illustrates operation of an embodiment of the present

25   invention on a router.

4

DETAILED DESCRIPTION OF THE EXEMPLARY EMBODIMENTS OF THE INVENTION

5   In the following Detailed Description of the Exemplary Embodiments of the

Invention, for purposes of explanation and not limitation, specific details are set forth in order

to provide a thorough understanding of embodiments of the present invention. However, it

will be apparent to those of ordinary skill in the art that embodiments of the present invention

can be practiced in other embodiments that depart from these specific details. In other

10   instances, detailed descriptions of well-known methods, devices, logical code (e.g.,

hardware, software, firmware), and the like are omitted so as not to obscure description of

embodiments of the present invention with unnecessary detail. In particular, aspects of the

BGP are referenced in order to describe aspects of embodiments of the present invention. It

should be understood by those having skill in the art that the present invention can be

practiced in embodiments that depart from the BGP.

        A system and method in accordance with principles of the present invention correlate

routing information base data from at least one router with corresponding data flow

information. The correlation is performed in order to compute data traffic volumes for a

plurality of autonomous system (AS) numbers. The system and method can aggregate and

20   calculate the traffic volumes of various network transit providers and then provide

information about how much network traffic transits or terminates at particular ASs. The

system includes at least one computer. The at least one computer collects data flow statistics

from routers that are configured to send data flow statistics to the at least one computer. The

computer then aggregates the data flow statistics. In preferred embodiments of the present

25   invention, clocks on each AS are synchronized to one another via a known time

synchronization protocol, such as, for example, the network timing protocol (NTP). Use of

5

such a known time synchronization protocol helps to ensure that measurement intervals used

by preferred embodiments of the present invention correspond to one another.

The data flow statistics are correlated with routing information base data by finding

which selected route in the routing information base data a given traffic flow traversed.

5   Using an AS path listed for a selected route, a counter is incremented by the size of the data

flow for each AS listed in the selected route. A set of counters, which represent data traffic

that transited or terminated at each AS, results. The counters can then be combined based on

network providers represented by each AS number. A report is created from the combined

counters. The report describes how much data traffic transited or terminated at a particular

10   provider's network.

Another computation can be performed by the at least one computer on a synthetic AS

path. The synthetic AS path is computed by combining all possible AS paths for a particular

prefix into a single path that retains the first and last AS and contains only one occurrence of

each intervening AS. In other words, the synthetic AS path is a union of all AS paths for a

15   prefix, with all duplicates removed. The effect of constructing the synthetic AS path in this

manner is that the best connected AS numbers will have high transit volumes. As such, they

will provide an indication to the network operator regarding the most attractive transit

providers and peers.

A report can be displayed by the at least one computer that shows, for a particular

20   time period, results of the two calculations. The report preferably indicates which of the

network transit providers is already a transit provider or a peer for the network under

evaluation. The report can be used to aid in the negotiation of contracts with current transit

providers based on documented and projected traffic volumes. The report can also be used to

choose more optimal transit providers or peers for the profile of data traffic the network

25   operator is handling. The report can profile traffic utilization of customers of the network

6

operator and can also perform calculations based on flows correlated with routing

information bases provided by potential transit providers or peers. Reports based on routing

information bases provided by potential transit providers or peers can be used by the network

operator to evaluate the effect of connecting to a new or different set of transit providers or

5   peers.

FIG. 1 shows an exemplary system 100 operating according to principles of the

present invention. The system 100 includes two routers R_A 102 and R_B 104. The routers R_A

102 and R_B 104 provide access to internet transit providers X 106, Y 108, and Z 110. The

system 100 also includes a data flow collection node N_F 112. The data flow collection node

10   N_F 112 collects data flow statistics (e.g., NETFLOW data) from the router R_A 102 and the

router R_B 104 during a sampling interval. The sampling interval can be variably set by an

administrator. The system 100 also includes a routing information base collection node N_B

114. The routing information base collection node N_B 114 collects at periodic intervals

routing information base data such as a BGP table from each of the router R_A 102 and the

15   router R_B 104. For example, the routing information base could be collected by the routing

information base collection node at the beginning of an hour-long sampling interval (i.e., a

snapshot at the beginning of the interval). A correlation node N_C 116 correlates the data

collected at the data flow collection node N_F 112 and the routing information base collection

node N_B 114. The correlation node N_C 116 correlates the data flow statistics with

20   corresponding routes traversed by the data flows to arrive at a determination of how much

data traversed each AS represented by a selected route of each of the router R_A 102 and the

router R_B 104. A reporting node N_R 118 analyzes and creates reports 120 on the correlated

data received from the correlation node N_C 116.

The routing information base collection N_B 114, the data flow collection node N_F 112,

25   the correlation node N_C 116, and the routing node N_R 118 can each be discrete nodes that are

7

external to one another. In the alternative, one or more of the routing information base

collection N_B 114, the data flow collection node N_F 112, the correlation node N_C 116, and the

routing node N_R 118 can be incorporated into a single node.

FIG. 2 is a flow chart that illustrates a flow 200 that describes how the correlation

5   node N_C 116 operates. The correlation node N_C 116 determines an AS path set of an

observed network data flow. The AS path set is the set of AS numbers appearing in a

selected AS path in the routing information base corresponding to the network data flow.

The correlation node N_C 116 preferably produces a table of AS transit volumes and a table of

AS terminating volumes. The AS terminating volume is the total volume of traffic

10   terminating at an AS for the entire routing table. An AS transit volume is the volume of data

traffic that transits a given AS for a selected AS path. A selected prefix AS path is the AS

path that was selected by the BGP from listed possible AS paths for the given prefix.

At step 202, the correlation node N_C 116 determines the routing information base

(e.g., BGP table) that corresponds to both of: 1) a particular router where a data flow was

15   observed; and 2) a sampling interval when the data flow was observed. At step 204, a data

flow destination address obtained from the data flow collection node N_F 112 is matched to a

prefix in the routing information base determined at step 202. Step 204 is preferably

performed by collecting a list of all prefixes in the routing information base determined at

step 202, ordering the list in ascending IP address order and descending netmask order, and

20   determining which prefix matches a given IP address. To determine which prefix matches

the given IP address, a search of the routing information base for a prefix that most closely

matches the given IP address is performed.

At step 206, the selected AS path that matches the prefix is extracted from the routing

information base. At step 207, a synthetic AS path is computed. A synthetic AS path is

25   computed by: 1) forming the union of all AS numbers for every AS path for a given prefix;

8

and 2) removing all duplicate AS numbers from the union. Embodiments of the present invention can, for example, use either synthetic AS paths, selected AS paths, or both. The synthetic AS path calculation is useful when a prefix has multiple path choices.

At step 208, the terminating AS number of the selected AS path is determined. At

5   step 210, a byte count for each AS number in the selected AS path is incremented by the data flow volume. In a preferred embodiment, steps 204-210 are repeated for each destination address obtained from the data flow collection node N_F 112.

FIG. 3 is a flow diagram that illustrates the operation of an embodiment of the present invention on a router. In contrast to the embodiment shown in FIG. 1, the flow 300 is

10   preferably performed on a router itself. The flow 300 begins at step 302, wherein data directed toward a destination address using a selected autonomous path in a routing information base of the router is detected. At step 304, a counter is incremented by an amount that indicates the volume of the data for each AS in at least one of the selected AS path and the synthetic AS path. The counter is maintained either on the router or is

15   transmitted to another entity external to the router.

In the event that one or more ASs in the routing information base of the router is updated, routing of further data causes the updated routing information base to be used for incrementing a counter corresponding to the ASs traversed by the further data. Therefore, updates to the routing information base are automatically accounted for as needed.

20   When analysis and reporting are needed for a plurality of routers operating according to the flow 300, counter data resulting from the incrementing of the counters for each of the plurality of routers during a specified time period can be obtained. The counter data is then analyzed and results of the analysis are reported. For example, counter data for the specified time period could be obtained and summed for each AS and then placed in reports 120 in

25   descending order.

9

Referring again to FIGS. 1 and 2, exemplary operation of the system 100 will now be described. Exemplary BGP tables for the router R$_A$ 102 and the router R$_B$ 104, respectively, are shown in Tables 1 and 2.

Table 1

R$_A$:	Prefix	Next Hop	AS Path	Selected
	9.20.0.0/16	157.22.9.37	70 24 2686	
		4.0.71.165	7018 24 2686	<<<<<
		166.63.50.121	3561 2555 2686	
	12.1.83.0/24	157.22.9.37	70 693 6461 14787	
		4.0.71.165	7018 24 14787	
		166.63.50.221	3561 24 14787	<<<<<

Table 2

R$_B$:	Prefix	Next Hop	AS Path	Selected
	9.20.0.0/16	157.24.9.12	70 24 2686	<<<<<
		4.0.72.164	7018 24 2686	
		166.63.40.111	3561 2555 2686	
	12.1.83.0/24	157.24.9.12	70 693 6461 14787	
		4.0.72.164	7018 24 14787	<<<<<
		166.63.40.111	3561 24 14787	

Each of Table 1 and Table 2 lists at least one prefix that is used to route data through the router R$_A$ 102 and R$_B$ 104, respectively. Each of the listed prefixes includes at least one next hop address and at least one AS path that are used to route data having a destination address matching the prefix. For each of the routers R$_A$ 102 and R$_B$ 104, one of the AS paths is a selected AS path. A selected AS path is used to route data having a destination address that matches the corresponding prefix.

The only major difference between the BGP table for the router R_A 102 (Table 1) and the BGP table for the router R_B 104 (Table 2) is which route is selected for a given prefix. Route selection is often based on policies that an administrator configures for a router and does not necessarily result in the shortest AS path. It is assumed that the BGP tables shown

5   in Tables 1 and 2 have been acquired by routing information base collection node N_B 114.

Table 3 shows exemplary data flow statistics collected by the data flow collection node N_F 112 from the router R_A 102. Table 4 shows exemplary data flow statistics collected by the data flow collection node N_F 112 from the router R_B 104. Both Table 3 and Table 4 list data flow volume by destination IP address.

Table 3 (R_A)

Destination	Size
9.20.0.1	200 MB
9.20.230.6	300 MB
9.20.130.37	600 MB
12.1.83.7	400 MB
12.1.83.26	500 MB

Table 4 (R_B)

Destination	Size
9.20.16.21	600 MB
9.20.66.36	600 MB
9.20.30.124	100 MB
12.1.83.3	400 MB
12.1.83.21	300 MB
12.1.83.72	200 MB

The data flow collection node N_F 112 has aggregated, per destination address, the data flow statistics it has collected during a sampling interval for each of the routers R_A 102 and

15   R_B 104. A program known as cflowd, which is a generally available software program, or

11

any other data flow collection program, can be used to aggregate the data flow statistics from

each of the routers R_A 102 and R_B 104 by destination address.

The correlation node N_C 116 retrieves both the aggregated data flow statistics from

the data flow collection node N_F 112 and the BGP tables for each of the routers R_A 102 and

5   R_B 104 from the routing information base collection node N_B 114. The correlation node N_C

116 then performs a correlation of the BGP tables and the aggregated data flow statistics to

yield data flow statistics. The data flow statistics are correlated for each AS in the selected

AS paths listed in the BGP tables collected by the routing information base collection node

N_B 114.

10   The correlation node N_C 116 performs the correlation by determining the BGP table

associated with the router R_A 102 as described in step 202. The determined BGP table for the

router R_A 102 is shown in Table 1. Steps 204, 206, 208, and 210 are then performed as

follows using the BGP table of the router R_A 102 and the data flow statistics found in Table

3:

9.20.0.1   200 MB   AS PATH 7018 24 2686

The data flow to 9.20.0.1 took the selected route from AS 7018 to AS 24 to AS 2686.

Since the data flow to 9.20.0.1 was 200 MB, C(7018) = 200MB, C(24) = 200MB, and

20   C(2686) = 200MB.

9.20.230.6   300 MB   AS PATH 7018 24 2686

The data flow to 9.20.230.6 took the selected route from AS 7018 to AS 24 to AS

5   2686, so C(7018) = C(7018) + 300MB = 500MB, C(24) = C(24) + 300MB = 500MB, and

C(2686) = C(2686) + 300MB = 500 MB.

9.20.130.37   600 MB   AS PATH 7018 24 2686

The data flow to 9.20.130.37 took the selected route from AS 7018 to AS 24 to AS

2686, so C(7018) = C(7018) + 600MB = 1100MB, C(24) = C(24) + 600MB = 1100MB, and

C(2686) = C(2686) + 600MB = 1100MB.

12.1.83.7   400 MB   AS PATH 3561 24 14787

The data flow to 12.1.83.7 took the selected route from AS 3561 to AS 14787, so

C(3561) = 400MB, C(24) = C(24) + 400MB = 1500MB, and C(14787) = 400MB.

12.1.83.26   500 MB   AS PATH 3561 24 14787

20

The data flow to 12.1.83.26 took the selected route from AS 3561 to AS 14787, so

C(3561) = C(3561) + 500MB = 900MB, C(24) = C(24) + 500MB = 2000MB, and C(14787)

= C(14787) + 500MB = 900MB.

At the conclusion of the correlation of the data flow statistics of Table 3 and the BGP

25   table of the router R_A 102 (Table 1), the AS counters are as follows:

13

C(70) = 0 MB

C(24) = 2000 MB

C(7018) = 1100 MB

5   C(2686) = 1100 MB

C(3561) = 900 MB

C(2555) = 0 MB

C(693) = 0 MB

C(6461) = 0 MB

10   C(14787) = 900 MB

The correlation node N_C 116 continues performing the correlation by determining the

BGP table associated with the router R_B 104 as described in step 202. The determined BGP

table for the router R_B 104 is shown in Table 2. Steps 204, 206, 208, and 210 are then

15   performed as follows using the BGP table of the router R_B 104 and the data flow statistics

found in Table 4:

9.20.16.21   600 MB   AS PATH 70 24 2686

20   The data flow to 9.20.16.21 took the selected route from AS 70 to AS 24 to AS 2686.

Since the data flow was 600 MB, C(70) = 600MB, C(24) = C(24) + 600MB = 2600MB, and

C(2686) = C(2686) + 600MB = 1700MB.

9.20.66.36   600 MB   AS PATH 70 24 2686

The data flow to 9.20.66.36 took the selected route from AS 70 to AS 24 to AS 2686,

so $C(70) = C(70) + 600MB = 1200MB$, $C(24) = C(24) + 600MB = 3200MB$, and $C(2686) =$ $C(2686) + 600MB = 2300MB$.

9.20.30.124   100 MB   AS PATH 70 24 2686

The data flow to 9.20.130.124 took the selected route from AS 70 to AS 24 to AS

2686, so $C(70) = C(70) + 100MB = 1300MB$, $C(24) = C(24) + 100MB = 3300MB$, and

$C(2686) = C(2686) + 100MB = 2400MB$.

12.1.83.3   400 MB   AS PATH 7018 24 14787

The data flow to 12.1.83.3 took the selected route from AS 7018 to AS 24 to AS

14787, so $C(7018) = C(7018) + 400MB = 1500MB$, $C(24) = C(24) + 400MB = 3700MB$, and

$C(14787) = C(14787) + 400MB = 1300MB$.

12.1.83.21   300 MB   AS PATH 7018 24 14787

The data flow to 12.1.83.21 took the selected route from AS 7018 to AS 24 to AS

14787, so $C(7018) = C(7018) + 300MB = 1800MB$, $C(24) = C(24) + 300MB = 4000MB$, and

$C(14787) = C(14787) + 300MB = 1600MB$.

12.1.83.72   200 MB   AS PATH 7018 24 14787

The data flow to 12.1.83.72 took the selected route from 7018 to AS 24 to AS 14787, so $C(7018) = C(7018) + 200MB = 2000MB$, $C(24) = C(24) + 200MB = 4200MB$, and

5     $C(14787) = C(14787) + 200MB = 1800MB$.

At the conclusion of the correlation for both the router R_A 102 and the router R_B 104, the AS data flow statistics are as follows:

C(70) = 1300MB

C(24) = 4200MB

C(7018) = 2000MB

C(2686) = 2400MB

C(3561) = 900MB

C(2555) = 0MB

C(693) = 0MB

C(6461) = 0MB

C(14787) = 1800MB

20   Referring again to FIG. 1, the overall AS data flow statistics are forwarded to the reporting node N_R 118. The overall AS data flow statistics can be made to indicate which AS numbers represent existing peers. For example, AS 70, AS 7018, and AS 3561 could be indicated as being existing peers.

The reporting node N_R 118 can prioritize the list of AS numbers according to their

25   overall data flow statistics as aggregated by the correlation node N_C 116 and indicate those

AS numbers that are existing peers. The reporting node N_R 118 can display, for example, the top 5 AS transit volumes as shown in Table 5, wherein a * indicates an AS number that is a current transit provider or peer.

Table 5

AS Number	Transit Volume
24	4200MB
2686	2400MB
*7018	2000MB
14787	1800MB
*70	1300MB

Table 5 shows that AS 24 would make an excellent peer because much traffic volume is routed through AS 24. There is much traffic traversing AS 2686. Therefore, AS 2686 would also make a good peer.

In another embodiment of the present invention, volume calculations are performed using synthetic AS paths as shown in step 207 of FIG. 2 in addition to or instead of the selected-route AS paths as described above. A synthetic AS path is computed by: 1) forming the union of all AS numbers for every AS path for a given prefix; and 2) removing all duplicate AS numbers from the union.

If the same routing tables listed in Tables 1 and 2 and the same data flow statistics listed in Tables 3 and 4 are used, Table 6 results. Table 6 includes each prefix and its associated synthetic AS path.

Table 6

Prefix	Synthetic AS Path
9.20.0.0/16	70 24 7018 3561 2555 2686
12.1.83.0/24	70 693 6461 7018 24 3561 14787

17

Each entry from Table 3 and 4 is processed as described above except that, for each

router, a synthetic AS path, as opposed to a selected AS path, is used. Each AS counter is

incremented by the flow sizes that match each prefix. The resulting volume calculations after

5   processing each entry are as follows:

C(70) = 4200

C(24) = 4200

C(7018) = 4200

C(3561) = 4200

C(2555) = 2400

C(2686) = 2400

C(693) = 900

C(6461) = 900

C(14787) = 900

The reporting node NR can report the top 5 synthetic transit volumes as shown in

Table 7, in which * indicates an AS number that is a current peer or transit provider:

20

Table 7

AS Number	Synthetic Transit Volume
*70	4200
24	4200
*7018	4200
*3561	4200
2555	2400

18

From this analysis, it is apparent that AS number 24 is a good potential transit provider.

The above illustrates how an exemplary system in accordance with principles of the present invention can operate. Although it would be easy to see from the BGP routing tables and data flow statistics that AS 24 would be a good peering candidate, real-world data sets are typically much more complicated and much more difficult to analyze. Embodiments of the present invention process data and yield reports 120 that permit a simplified view of what AS numbers are most active and, thus, which AS numbers would be best to peer with for routing transit traffic.

As will be recognized by those having skill in the art, the innovative concepts described in the present patent application can be modified and varied over a wide range of applications. Accordingly, the scope of patented subject matter should not be limited to any of the specific exemplary teachings discussed, but is instead defined by the following claims.